PARAMETER ESTIMATION FOR A TWO-STATE SEMI-MARKOV MODEL OF A UNIVARIATE POINT PROCESS

James Leroy Hornback

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

PARAMETER ESTIMATION FOR A TWO-STATE
SEMI-MARKOV MODEL OF A
UNIVARIATE POINT PROCESS

James Leroy Hornback

Thesis Advisor: P. A. W. Lewis

March 1974

Parameter Estimation for a Two-State

Semi-Markov Model of a Univariate Point Process

by

James Leroy Hornback
Lieutenant, United States Navy
B.A., San Fernando Valley State College, 1968

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1974

ABSTRACT

Using the convenient second-order interval properties of a two-state semi-Markov model for a univariate point process, an automated technique for the estimation of the parameters in the model was researched and discussed. The power spectral density of intervals was estimated by the periodogram and a Kolmogorov-Smirnov test of fit was conducted. The asymtotic exponential distribution and independence of the periodogram points were used to calculate an approximate likelihood function. A system of equations was then formed to find the maximum likelihood estimates of the parameters. Since closed-form solutions for the estimates could not be found, an iterative method to stabilize initial guesses of the parameter values was attempted with only limited success. Results on using Kolmogorov-Smirnov type statistics and the spectrum of intervals to test the fit of stochastic process models to data have also been obtained.

TABLE OF CONTENTS

I. INTRODUCTION -- 5

II. BACKGROUND ANALYSIS -------------------------------- 8

A. TWO-STATE SEMI-MARKOV MODEL FOR UNIVARIATE POINT PROCESS ---------------------------------- 8

B. PARAMETER ESTIMATION TECHNIQUE --------------- 10

III. EXPERIMENTAL APPROACH ---------------------------- 15

A. UNIVARIATE TWO-STATE SEMI-MARKOV SIMULATION -- 15

B. CALCULATION OF THE PERIODOGRAM --------------- 17

C. SOLVING SIMULTANEOUS EQUATIONS FOR THE PARAMETERS ------------------------------------ 20

IV. RESULTS AND CONCLUSIONS -------------------------- 24

COMPUTER SUBPROGRAMS ----------------------------------- 29

LIST OF REFERENCES ------------------------------------- 38

INITIAL DISTRIBUTION LIST ------------------------------ 39

FORM DD 1473 --- 40

ACKNOWLEDGEMENT

The author gratefully acknowledges the cooperation and assistance of the IBM San Jose Research Laboratory which generously provided computer research facilities during a curriculum sponsored, six week, experience tour. In particular, the support and guidance of Mr. G. S. Shedler, IBM San Jose, and Professor P. A. W. Lewis, Naval Postgraduate School, were greatly appreciated.

I. INTRODUCTION

It is in the nature of the Operations Research approach to the study of problems to attempt the construction of a mathematical model for the problem. Subclasses of mathematical models include stochastic, i.e. utilizing random variables, and deterministic models. If a stochastic model seems appropriate and a general model is proposed, it remains necessary to estimate parameters of the model from data. Parameter estimates, as well as the general form of the model, usually come from detailed analysis of data observed from the problem or process under investigation.

Several techniques utilizing observed data exist for the estimation of parameters for stochastic models. Typically the methods of moments or maximum likelihood are used and usually yield estimates with some desirable properties. Methods such as these frequently require the simultaneous solutions to a system of equations in order to find estimates. A number of computer approximation routines have been developed for the solution of such systems, but their usefulness seems limited.

One proposed stochastic model provided the impetus for this research. Lewis and Shedler [1973], while studying page reference patterns in a demand paged computer system, formulated a univariate two-state semi-Markov model for the process of page exceptions. Page exceptions occur because a

computer program which is in execution has been stored in blocks of storage called pages. Some of these pages must be in core storage for the program to be executing, while the remaining pages may be located on peripheral storage devices. Following the execution of each instruction a page is referenced which contains the next instruction. If this referenced page is in core storage execution continues; however, if the referenced page is not in core storage then execution is interrupted and the referenced page must be read into core storage. This type of interruption is referred to as a page exception. Data for this process was generated by counting the number of page references occurring between page exceptions. Lewis and Shedler [1973] discussed their procedure for estimating parameters which they described as an ad hoc method, and concluded that there was a need to formalize the parameter estimation procedure.

The purpose of the research in this thesis was to utilize the convenient second-order interval properties of a univariate two-state semi-Markov process to produce an automated, computer programed, technique for the estimation of parameters for the model. This was desirable because the ad hoc method used by Lewis and Shedler [1973] was very time-consuming and there exists a considerable body of page exception data which it is desired to analyze. The basic procedure was to calculate an estimate for the power spectral density of the process, namely the periodogram, and

utilize an approximate method of maximum likelihood to estimate the parameters.

It will be seen that the proposed procedure did not work as well as hoped, but the problems which arose pointed up other possible attacks on the problem. It should also be noted that model fitting and parameter estimation for these point processes is almost a completely open field.

II. BACKGROUND ANALYSIS

A. TWO-STATE SEMI-MARKOV MODEL FOR UNIVARIATE POINT PROCESS

Excellent discussions of this model can be found in Cox and Lewis [1966,Ch.7] and Lewis and Shedler [1973]. Those discussions are summarized here for continuity of exposition.

Let the sequence of random variables $\{X_i, i=1,\ldots,N\}$ be interevent times, i.e. X_i is the interevent time between event (i-1) and event (i). In order that a discussion of equilibrium distributions may be avoided it was assumed that a hypothetical event has occurred at time zero, so that X_1, the interval between time zero and the first event, is an observation from the same process as the remainder of the sequence, i.e. there is no length-biased sampling [Cox and Lewis 1966,Ch.4] included.

Now suppose there are two types of intervals but that the interval type is not observable, i.e. a univariate point process. The two interval types have probability mass functions (p.m.f.) $p_1(x)$ and $p_2(x)$, respectively, with transitions between types described by a two-state Markov chain with matrix

$$\underline{A} = \begin{pmatrix} \alpha_1 & 1-\alpha_1 \\ 1-\alpha_2 & \alpha_2 \end{pmatrix}$$

That is, given that X_i has p.m.f. $p_2(x)$ then X_{i+1} has p.m.f. $p_2(x)$ with probability α_2 and p.m.f. $p_1(x)$ with probability $1-\alpha_2$, independent of the history of previous intervals, etc.

The vector of steady-state probabilities $\underline{\pi} = (\pi_1 \ \pi_2)$ associated with the transition matrix $\underline{A}$ results from the solution of the matrix equation $\underline{\pi} = \underline{\pi A}$ and it follows that

$$\pi_1 = \frac{1-\alpha_2}{2-\alpha_1-\alpha_2} \qquad \pi_2 = \frac{1-\alpha_1}{2-\alpha_1-\alpha_2} \quad .$$

If μ_1, σ_1^2 and μ_2, σ_2^2 are the mean and variance for intervals with p.m.f. $p_1(x)$ and $p_2(x)$, respectively, the steady-state marginal results for intervals between events in the univariate process, i.e. interval type not known, are as follows:

$$p(x) = \pi_1 p_1(x) + \pi_2 p_2(x) \ ,$$

$$\mu = E(X) = \pi_1 \mu_1 + \pi_2 \mu_2 \ ,$$

$$\sigma^2 = \text{var}(X) = \pi_1 \sigma_1^2 + \pi_2 \sigma_2^2 + \pi_1 \pi_2 (\mu_1 - \mu_2)^2 .$$

The serial correlation coefficients of lag k, ρ_k, for the intervals are of the form $m\beta^k/\sigma^2$ where, for $k=1,2,\ldots,$

$$m = \pi_1 \pi_2 (\mu_1 - \mu_2)^2 \qquad \beta = \alpha_1 + \alpha_2 - 1 \ .$$

From these coefficients the positive portion of the power spectral density may be computed,

$$P_+(\omega_n) = \frac{\sigma^2}{\pi} \left(1 + 2 \sum_{k=1}^{\infty} \rho_k \cos k\omega_n\right).$$

The closed-form solution to the infinite series is given by Jolley [1961, series #545] yielding

$$P_+(\omega_n) = \frac{\sigma^2}{\pi} \left[1 + 2 \frac{m\beta}{\sigma^2} \left\{\frac{(\cos \omega_n) - \beta}{1+\beta^2-2\beta\cos \omega_n}\right\}\right]. \qquad (1)$$

The beneficial feature of the power spectral density for this model is that it only depends upon the mean and variance of each of the two probability distributions and not on the complete distributions, and is thus fairly robust.

The count spectrum [Cox and Lewis 1966, Ch.4] on the other hand, depends on the complete distributions.

B. PARAMETER ESTIMATION TECHNIQUE

Lewis and Shedler [1973] used a modified method of moments approach in order to estimate the parameters in their model. The standard method of moments procedure for parameter estimation is to calculate theoretical moments in terms of the unknown parameters and equate them to empirical observations of these moments. An alternative to this method is the method of maximum likelihood. In this method parameter values are selected which maximize the joint probability density of the observed data. To accomplish this there is a need for some distributional assumptions. However, even for a simple model such as the univariate two-state semi-Markov model discussed here it is not possible to write down the joint density of the intervals. Thus the following approximate technique was proposed and tried.

It is known, Cox and Lewis [1966,Ch.5], that an estimate of the power spectral density $P(\omega_n)$ at ω_n, the periodogram $I(n)$, is in general asymtotically exponentially distributed [Olshen 1967]. The periodogram is an unbiased estimate, i.e. $E[I(n)]=P(\omega_n)$; however, it is not a consistent estimate since the variance of the exponential distribution is equal to the square of the mean, i.e. the variance does not decrease with increased sample sizes. Moreover, for n_1 not equal to n_2 the periodogram points $I(n_1)$ and $I(n_2)$ are

asymtotically independent. Thus for finite sample size N an approximate likelihood function may be written by assuming the periodogram points are independent with exponential distributions having mean value $P(\omega_n)$. This is the technique explored in this thesis.

The definition and development of the periodogram requires the finite Fourier transform.

The finite Fourier transform was discussed by Cooley, Lewis and Welch [to be published in 1974]. Let $\{Y(j), j=0,\ldots,N-1\}$ be a sequence of N real numbers. The finite Fourier transform of Y(j) is then

$$a(n) = \frac{1}{N}\sum_{j=0}^{N-1} Y(j)e^{-\frac{2\pi inj}{N}}, \quad n=0,\ldots,N-1.$$

This sequence of complex numbers may also be written in the form

$$a(n) = \frac{1}{N}\left[\sum_{j=0}^{N-1} Y(j)\cos(j\omega_n) - i\sum_{j=0}^{N-1} Y(j)\sin(j\omega_n)\right],$$

where $\omega_n = \frac{2\pi n}{N}$, $n=0,1,\ldots,N-1$. The periodogram I(n) is then

$$I(n) = \frac{N|a(n)|^2}{2\pi}, \quad n=0,\ldots,$$

or

$$I(n) = \frac{\{\sum_{j=0}^{N-1} Y(j)\cos(j\omega_n)\}^2 + \{\sum_{j=0}^{N-1} Y(j)\sin(j\omega_n)\}^2}{2\pi N}.$$

The periodogram is an even, periodic, function and hence has only [N/2]+1 distinct values, where [N/2] is the integer part of N/2. Hereafter in the discussion N will refer to an even integer. Let $I_+(n)=2I(n)$ be the estimate for $P_+(\omega_n)$.

It is easily seen that $I_+(0)$ is proportional to the square of the arithmetic average of the observed data; thus no new information is obtained from $I_+(0)$. Since N/2 is an integer $\omega_{N/2}=\pi$ and $I_+(N/2)$ is proportional to the square of an alternating summation of the data. Both $I_+(0)$ and $I_+(N/2)$ were ignored in what follows, thus leaving (N/2)-1 periodogram points. It should be added that these two periodogram points, suitably normalized, have asymtotic χ^2 distributions with one degree of freedom and not an exponential distribution.

Now there is sufficient information to begin the approximate maximum likelihood search for parameter estimates. The parameters of this model that need estimation are the mean and variance of each marginal distribution and the two transition probabilities α_1 and α_2. As a vector these parameters will be labeled $\underline{\theta}=(\mu_1,\sigma_1{}^2,\mu_2,\sigma_2{}^2,\alpha_1,\alpha_2)$ and individually, to simplify notation, as θ_j, j=1,2,3,4,5,6, to stand for the parameter as an element of the vector $\underline{\theta}$.

The approximate likelihood function can be written as

$$L(\underline{\theta}) = \prod_{n=1}^{(N/2)-1} \frac{1}{P_+(\omega_n;\underline{\theta})} e^{-I_+(n)/P_+(\omega_n;\underline{\theta})},$$

which is equivalent to

$$L(\underline{\theta}) = \left(\prod_{n=1}^{(N/2)-1} \frac{1}{P_+(\omega_n;\underline{\theta})}\right) e^{-\sum_{n=1}^{(N/2)-1} \frac{I_+(n)}{P_+(\omega_n;\underline{\theta})}}.$$

A more simple function to work with, which has the same maximum as $L(\underline{\theta})$, is the log likelihood function $LL(\underline{\theta})=\ln L(\underline{\theta})$, where ln symbolizes the natural logarithm.

Then

$$LL(\underline{\theta}) = -\sum_{n=1}^{(N/2)-1} \ln\{P_+(\omega_n;\underline{\theta})\} - \sum_{n=1}^{(N/2)-1} \frac{I_+(n)}{P_+(\omega_n;\underline{\theta})} .$$

In the typical mathematical approach to finding an unconstrained maximum of a function, it is a necessary condition that all of the first-order partial derivatives of the function, with respect to the unknown parameters, be equal to zero, i.e. that

$$0 = \frac{\partial LL(\underline{\theta})}{\partial \theta_j} = LL_j = \sum_{n=1}^{(N/2)-1} \frac{I_+(n)-P_+(\omega_n;\underline{\theta})}{P_+(\omega_n;\underline{\theta})} P_j, j=1,\ldots,6 \qquad (2)$$

where P_j and LL_j are the first-order partial derivatives of $P_+(\omega_n;\underline{\theta})$ and $LL(\underline{\theta})$, respectively, with respect to parameter θ_j. This process results in six equations and six unknown parameters. Parameter estimates are found by simultaneously solving the system of equations for each of the parameters, although this may not yield a unique maximum. If the system is of a simple form it may be possible to get at least a few closed-form solutions which will reduce the size of the system.

Once the parameter estimates have been found it is necessary to show that a maximum has been achieved. A sufficient condition for a maximum is that the matrix of second-order partial derivatives be negative definite. The final phase in this approximate likelihood estimation process is to verify the predictability of the model. The verification may be done, using the estimated parameters, by calculating other theoretical properties of the model, such as the spectrum of counts discussed by Cox and Lewis [1966, Ch.4],

which may then be compared with the corresponding empirical properties of the data. Note that the utility of the spectrum of intervals in the approximate likelihood analysis is that it does not depend on the complete distributional form for $p_1(x)$ and $p_2(x)$ while the spectrum of counts does. It will be seen later, however, that this independence leads to ill-conditioning in the solution of the maximum likelihood equations.

III. EXPERIMENTAL APPROACH

The original data, analyzed by Lewis and Shedler [1973], was not available for this research. In view of this fact and since the purpose of the research was to evaluate the effectiveness of the previously described technique for parameter estimation it was felt that data observed from a model with known parameters would better aid the evaluation process. With this in mind a simulation of the model described by Lewis and Shedler [1973] was constructed for the purpose of generating such data.

A. UNIVARIATE TWO-STATE SEMI-MARKOV SIMULATION

The simulation, as well as the model, was subdivided into three major subsections. The state transition matrix $\underline{A}$ was one subsection and the two distributions for intervals were the remaining two subsections. Lewis and Shedler [1973] postulated a geometric distribution for the long intervals and a negative binomial distribution for the shorter intervals. The parameters used for the simulation were those calculated by Lewis and Shedler [1973].

A Monte Carlo simulation, such as this, required a pseudo-random number generator with favorable serial correlation properties. Learmonth and Lewis [1973] discussed such a generator called SRAND. SRAND returns an observation from a standardized uniform distribution on the interval

(0,1). SRAND is a multiplicative generator with a multiplier of (7^5) and a modulus of ($2^{31}-1$).

The geometric distribution is of the form

$$p_1(x) = p_1^{x-1}\ (1-p_1),\ 0<p_1<1;\ x=1,2,\ldots,$$

with a mean $\mu_1=1/(1-p_1)$ and variance $\sigma_1{}^2=p_1/(1-p_1)^2$. Utilizing the survivor function of the geometric distribution, i.e. $\text{prob}\{X>x\}=p_1{}^x$, $x=1,2,\ldots$, a generator of geometric variates was obtained. It was of the form

$$x = \lceil\{\ln(R)/\ln(p_1)\}\ ,$$

where R was an observation from SRAND and the symbol $\lceil\{b\}$ signified the smallest integer greater than or equal to b.

The negative binomial distribution is of the form

$$p_2(x) = \binom{k+x-2}{x-1} p_2{}^{x-1}\ (1-p_2)^k\ ,$$

$0<p_2<1$, $k>0$, $x=1,2,\ldots$, with mean $\mu_2=1+\{kp_2/(1-p_2)\}$ and variance $\sigma_2{}^2=kp_2/(1-p_2)^2$. Let $X|\lambda$, denoting X given a fixed value of λ, be distributed as a Poisson random variable with parameter λ. Now let λ have a gamma distribution with parameters k and η,

$$f(\lambda) = \frac{\eta^k\ \lambda^{k-1}\ e^{-\eta\lambda}}{\Gamma(k)},\ k>0;\ \lambda>0;\ \eta>0.$$

It can be shown using generating functions that the unconditioned X has a negative binomial distribution with parameters k and $p_2=1/(1+\eta)$.

To calculate a gamma variate with a parameter k, a positive real number, it was necessary to employ Johnk's technique [1964] for generating variates with the fractional

part of k. Let $\underline{k}$ be the integer part of k, if $k \geq 1$, or zero if $k<1$, and let $\bar{k}$ be the fractional part of k. The sum, λ_1, of $\underline{k}$ exponentially distributed random variables with parameter η has a gamma distribution with parameters $\underline{k}$ and η. In Johnk's technique let U_1, U_2 and U_3 be independent and identically distributed observations from a uniform distribution on interval (0,1) such that

$$Y = U_1^{1/\bar{k}} + U_2^{1/(1-\bar{k})} < 1 .$$

If $Y \geq 1$ new observations for U_1 and U_2 should be obtained. Then for $Z=U_1^{1/\bar{k}}/Y$ and $E=-\ln U_3$, $\lambda_2=(Z \times E)/\eta$ has a gamma distribution with parameters $\bar{k}$ and η. Finally $\lambda=\lambda_1+\lambda_2$ has the required gamma distribution with parameters k and η.

The generation of Poisson random variates with parameter λ was accomplished by letting X be equal to the largest integer n such that, for a sequence of independent identically distributed uniform random variates (U_i) from the interval (0,1),

$$U_1 \times U_2 \times \ldots \times U_n > e^{-\lambda} .$$

If $U_1 \leq e^{-\lambda}$ then $X=0$. X is then distributed as a Poisson variate with parameter λ.

B. CALCULATION OF THE PERIODOGRAM

The finite Fourier transform discussed in section II.B above requires on the order of N^2 complex operation pairs, i.e. a multiplication and an addition. For large N this can be very costly in terms of calculation time. Cooley, Lewis and Welch [1970] discussed the use of a fast Fourier

transform algorithm which only requires on the order of $N(r_1+\ldots+r_m)$ complex operation pairs where $N=(r_1\times\ldots\times r_m)$, i.e. r_m is a factor of N. The International Mathematical and Statistical Library [1973 revision] contains a computer subroutine, FFTR, which computes the fast Fourier transform of a real data sequence. For N=820, as in this research, the fast Fourier transform algorithm used only six percent of the number of complex operation pairs required by the straight-forward calculation method. Thus a significant savings in computer operating time was realized.

Utilizing previously described equations the periodogram $I_+(n)$ was computed and then used in a test of fit to the power spectral density $P_+(\omega_n)$. Cox and Lewis [1966, Ch.6] described a test based on the uniform distribution. While the periodogram has, asymtotically, an exponential distribution with mean $P_+(\omega_n)$, the quantity $I_+(n)/P_+(\omega_n)$ has an exponential distribution with mean one. This is true for each of the (N/2)-1 periodogram points. If all (N/2)-1 of these quantities are summed the total gives an interval of length over which there are (N/2)-2 points dispersed. The intervals between these points are each, hypothetically, an observation from a unit exponential distribution, i.e. the points form a Poisson process. It is a well known fact of the Poisson process that given M points are in an interval the M points are dispersed uniformly over the interval.

Thus, the sequence $\{U_{(i)}, i=1,...,(N/2)-2\}$, where

$$U_{(i)} = \frac{\sum_{n=1}^{i} I_+(n)/P_+(\omega_n)}{\sum_{n=1}^{(N/2)-1} I_+(n)/P_+(\omega_n)}$$

are uniform order statistics. The empirical cumulative distribution function for these quantities was then compared with the uniform cumulative distribution function using the Kolmogorov-Smirnov test. The null hypothesis is that the sequence $\{U_{(i)}\}$ is formed of uniform order statistics, while the alternative hypothesis remains unspecified. Lilliefors [1969] found that the critical values of the K-S test are too conservative when testing using exponential distributions where the mean has been estimated, as in this case. Too conservative means that the listed critical value for a level of significance α has actually a level of significance less than α. If the above test, with modified percentage points, accepts the null hypothesis then the assumption of a semi-Markov model for the data has been justified.

In order to test the periodogram it was necessary to know $P_+(\omega_n)$. As discussed earlier the correlation coefficient of lag k, ρ_k, is $\rho_k = m\beta^k/\sigma^2$ for this model. Let $\tilde{\gamma}(0)$, $\tilde{\gamma}(1)$ and $\tilde{\gamma}(2)$ be estimates of the variance and covariances of lags one and two, respectively, for the intervals. Then

$$\gamma(1) = \sigma^2\rho_1 = m\beta$$

and

$$\gamma(2) = \sigma^2\rho_2 = m\beta^2 \quad .$$

Solving simultaneously for $\tilde{m}$ and $\tilde{\beta}$, the estimates of m and β are

$$\tilde{\beta} = \frac{\tilde{\gamma}(2)}{\tilde{\gamma}(1)} \qquad \text{and} \qquad \tilde{m} = \frac{\tilde{\gamma}^2(1)}{\tilde{\gamma}(2)} \quad .$$

From (1) an estimate for $P_+(\omega_n)$ was

$$\tilde{p}_+(\omega_n) = \frac{1}{\pi} \{\tilde{\gamma}(0) + 2\tilde{m}\tilde{\beta} \frac{(\cos \omega_n) - \tilde{\beta}}{1+\tilde{\beta}^2-2\tilde{\beta} \cos \omega_n}\} \quad .$$

These estimates were then used in the computations for the sequence $\{U_{(i)}\}$.

C. SOLVING SIMULTANEOUS EQUATIONS FOR THE PARAMETERS

The system of equations defined by (2) and (1) was extremely complex, with no hope of finding a closed-form solution for any of the parameters. The system was reduced, however, by noting from the geometric distribution assumption that the variance for the long intervals was a function only of the parameter p_1, which also was the only parameter in the mean. It was a simple matter then to find the variance as a function of the mean which then reduced the system to only five unknown parameters. The system was still complex and required some iterative method for solution.

Rao [1965] suggested an iterative method which he called the Method of Scoring. He called LL_j the jth efficient score. The approach for this method was to assume some initial trial solution. Using a first-degree Taylor's expansion of the efficient scores about the trial solution, a system of linear equations was derived from which an additive correction to the trial solution was found. The

iterations were repeated until the additive corrections became negligible.

Specifically, let $\theta_1^0, \ldots, \theta_5^0$ be the trial values for the unknown parameters. From (2)

$$LL_j \simeq \frac{\partial LL(\underline{\theta})}{\partial \theta_j^0} + \sum_{i=1}^{5} (\theta_i - \theta_i^0) \frac{\partial^2 LL(\underline{\theta})}{\partial \theta_j^0 \partial \theta_i^0} , \; j=1,\ldots,5 \quad ,$$

where $\partial LL(\underline{\theta})/\partial \theta_j^0 = S_j^0$ the first-order partial derivative of $LL(\underline{\theta})$ with respect to θ_j, evaluated at θ_j^0. Let $\delta\theta_j = (\theta_j - \theta_j^0)$ and also let $\partial^2 LL(\underline{\theta})/(\partial \theta_j^0 \partial \theta_i^0) = T_{ji}^0$. Then

$$-\sum_{i=1}^{5} T_{ji}^0 \delta\theta_i = S_j^0 , \; j=1,\ldots,5$$

was a system of linear equations with five unknowns. In matrix notation the system had the form $\underline{-T\delta\theta} = \underline{S}$. Finally, the additive corrections were obtained from the equation $\underline{\delta\theta} = -\underline{T}^{-1}\underline{S}$ where $\underline{T}^{-1}$ was the inverse of the matrix $\underline{T}$, assuming $\underline{T}$ was nonsingular. The new trial solution then became $\underline{\theta}^1 = \underline{\theta}^0 + \underline{\delta\theta}$.

Rao [1965] explained that the variance of the final estimate θ_j^f of θ_j was approximated by the jth diagonal element of the matrix $(-\underline{T}^{-1})$. Recalling that the matrix of second-order partial derivatives of $LL(\underline{\theta})$ should be negative definite, then $-\underline{T}$ and $(-\underline{T}^{-1})$ were both positive definite.

In order to apply the method of scoring it was necessary to determine initial estimates of the parameters. The mean and variance of the intervals and the parameters m and β all have been estimated. Utilizing the marginal properties of the model and the method of moments a system of four

equations was developed which was of the form

$$\overline{X} = \frac{(1-\tilde{\alpha}_2)\tilde{\mu}_1 + (1-\tilde{\alpha}_1)\tilde{\mu}_2}{(1-\tilde{\beta})} \quad ;$$

$$\tilde{\gamma}(0) = \frac{(1-\tilde{\alpha}_2)(\tilde{\mu}_1{}^2-\tilde{\mu}_1) + (1-\tilde{\alpha}_1)\tilde{\sigma}_2{}^2}{(1-\tilde{\beta})} + \tilde{m} \quad ;$$

$$\tilde{\beta} = \tilde{\alpha}_1+\tilde{\alpha}_2-1 \quad ;$$

$$\tilde{m} = \frac{(1-\tilde{\alpha}_1)(1-\tilde{\alpha}_2)(\tilde{\mu}_1-\tilde{\mu}_2)^2}{(1-\tilde{\beta})^2} \quad ,$$

where $\overline{X}$ was the estimate for the mean of the intervals and the quantity $(\tilde{\mu}_1{}^2-\tilde{\mu}_1)$ was the estimate for $\sigma_1{}^2$ after making the geometric distribution assumption. From this system initial estimates for four parameters, as functions of the fifth parameter, were found and had the form

$$\tilde{\mu}_2 = \frac{\overline{X}\tilde{\mu}_1 - \overline{X}^2 - \tilde{m}}{(\tilde{\mu}_1-\overline{X})} \quad ;$$

$$\tilde{\alpha}_1 = \frac{(1-\tilde{\beta})\overline{X} - (\tilde{\mu}_2-\tilde{\beta}\tilde{\mu}_1)}{(\tilde{\mu}_1-\tilde{\mu}_2)} \quad ;$$

$$\tilde{\alpha}_2 = \tilde{\beta}+1-\tilde{\alpha}_1 \quad ;$$

$$\tilde{\sigma}_2{}^2 = \frac{(1-\tilde{\beta})(\tilde{\gamma}(0)-\tilde{m}) - (1-\tilde{\alpha}_2)(\tilde{\mu}_1{}^2-\tilde{\mu}_1)}{(1-\tilde{\alpha}_1)} \quad .$$

It only remained then to estimate parameter μ_1.

Lewis and Shedler [1973] explained that their estimate of μ_1 involved an eyeball judgement of where linearity began in the tail of the log survivor function of the data. This linearity in the tail led to the postulate of the geometric distribution for the long intervals. Since only an initial estimate was needed, their method of estimating μ_1 was utilized again. The value of the interval where linearity

began was subtracted from each greater interval and the arithmetic average of these intervals was then taken as the estimate of μ_1. Now all parameters had been initially estimated and the iterative method of scoring was applied to stabilize these estimates.

IV. RESULTS AND CONCLUSIONS

As the research for this thesis progressed two areas developed results which need to be discussed. The first of these was the test for justification of the exponential distribution assumption for the periodogram points. A subroutine called KSTEST was written to conduct this test. Part of the output of this subroutine was the Kolmogorov-Smirnov statistic which then was compared to a critical value from the distribution proposed by Lilliefors [1969] for the case where the mean of the exponential distribution had to be estimated. The 0.99 quantile of that distribution, i.e. a one percent level of significance, was 1.25. It was noted that, at this level, of four thousand trials made approximately six percent were rejected as not having produced periodograms from a semi-Markov model.

In addition to testing the hypothesis for each simulation another benefit was received. Since the testing did not strictly conform to that discussed by Lilliefors [1969], because each periodogram point had a different mean, it was felt that, for this case, quantiles of the distribution should be estimated. The four thousand data points of the statistic were obtained from four computer runs each containing one thousand simulations. For each run, the data was divided into ten sections in serial order, i.e. the first one hundred points were the first section, etc. The elements of each section were ordered and the 0.80, 0.85, 0.90,

0.95 and 0.99 empirical quantiles were observed. This resulted in ten observations for each quantile from which a mean and variance were estimated. Lastly, the entire data for the run was ordered and the five quantiles were observed. Thus, for each run, each of the five empirical quantiles had been observed and had an estimated mean and variance. Finally a mean of the four overall observations for each quantile and a mean of the section means were computed. The results are shown in Table I.

Lilliefors [1967] discussed the Kolmogorov-Smirnov test for normal data and calculated numerically the quantiles of the test statistic for the case where the mean and variance of the normal distribution must be estimated. These quantiles are included for comparison.

The second of the two significant areas was the estimation of parameters. The subroutine ESTIM8 was written, in double precision, to utilize the method of scoring for parameter estimation. As a result of the use of the subroutine several potential hazards to the proposed technique became visible.

The first of these hazards was the disparity between magnitudes of the five unknown parameters. Three of these are means and variances and the other two are probabilities, which are always less than or equal to one. This problem became apparent when the magnitudes of the scores and the elements in the matrix of second-order partial derivatives were seen. An attempt to correct this problem was made by

TABLE I

Source	Level of Significance				
	0.20	0.15	0.10	0.05	0.01
Usual quantiles	1.07	1.14	1.22	1.36	1.63
Lilliefors quantiles	0.86	0.91	0.96	1.06	1.25
Run 1	0.74	0.79	0.96	1.66	8.56
Mean	0.73	0.80	0.92	1.59	6.04
Variance	0.002	0.005	0.01	0.31	13.24
Run 2	0.75	0.85	1.03	2.05	6.66
Mean	0.75	0.85	1.03	1.86	7.11
Variance	0.003	0.01	0.04	0.45	25.35
Run 3	0.74	0.80	0.90	1.31	6.79
Mean	0.73	0.79	0.91	1.80	4.90
Variance	0.001	0.002	0.01	1.31	9.87
Run 4	0.73	0.82	0.96	1.72	8.45
Mean	0.74	0.83	0.95	1.64	5.90
Variance	0.01	0.02	0.03	0.62	12.51
Mean of Runs	0.74	0.82	0.96	1.69	7.62
Variance of Runs	0.00	0.00	0.001	0.03	0.27
Mean of Means	0.74	0.82	0.95	1.72	5.99
Variance of Means	0.00	0.00	0.001	0.01	0.21
Lilliefors normal quantiles	0.736	0.768	0.805	0.886	1.031

dividing the three large parameters by the overall mean of the intervals. This also favorably affected the partial derivatives involving these parameters. The desired effect was achieved in that the gap between magnitudes was narrowed; however, the parameter estimates that resulted from this modification were only about one percent different from the parameters achieved earlier, so apparently the disparity created no significant problem.

The second of these problems was that the final parameter estimates, overall, appeared to have little relationship to the marginal parameter values from which the data was generated. Similarly, parameter estimates for two sets of data differed greatly in magnitude and at times in sign, even when the periodogram was accepted as a close fit to the power spectral density. Differences in sign were extremely disturbing since all of the parameters were expected to be greater than zero.

A third problem, related to the second, was that the results failed, numerically, to establish that the matrix of second-order partial derivatives was negative definite. Similarly the negative inverse of that matrix could not be shown to be positive definite. This problem indicated that either a maximum had not been achieved, even though a cut-off criterion of 10^{-10} was used to test for convergence, or that due to round-off error the properties of a maximum could not be detected. With a smaller cut-off criterion the process would not converge and had to be terminated.

Finally, in a few instances when four of the five unknown parameters appeared close to the simulation parameters the initial value of the fifth parameter was changed and the subroutine was restarted. The parameters would again converge; however, the final values in some cases changed drastically, even to the point of changing sign.

Some of these problems may have been caused by an ill-conditioned system of equations, while others might be due to the lack of a powerful iterative technique for the solution of a system of equations that has, perhaps, poor initial estimates. In any case it should be clear that the use of second-order properties of a model might simplify or at least aid the parameter estimation process. One proposed modification to the technique discussed in this thesis was to use a mixture of the method of moments approach on the marginal distribution of the intervals and the maximum likelihood approach on the second-order properties to estimate parameters.

In conclusion it should be recalled that model fitting and parameter estimation for univariate point processes is almost a completely open field and that attempts, even unsuccessful ones, are needed in order to break-through the barrier of inadequate methodology.

```
C                          SUBROUTINE MODEL                              C
C                                                                        C
C     A.   PURPOSE:                                                      C
C          THIS SUBROUTINE CONTROLS THE SIMULATION OF A                  C
C          UNIVARIATE TWO-STATE SEMI-MARKCV POINT PROCESS.               C
C          THE STATE ONE INTERVALS HAVE A GECMETRIC                      C
C          DISTRIBUTION AND THE STATE TWO INTERVALS HAVE                 C
C          A NEGATIVE BINOMIAL DISTRIBUTICN.                             C
C                                                                        C
C     B.   USAGE:                                                        C
C          1.  ARGUMENTS:                                                C
C                                                                        C
C              X - OUTPUT VECTOR OF INTEREVENT TIMES (REAL*8)            C
C                                                                        C
C              SIZE - INPUT LENGTH OF VECTOR X (INTEGER)                 C
C                                                                        C
C              IX - INPUT RANDOM NUMBER SEED (INTEGER)                   C
C                                                                        C
C              DIST1M - INPUT MEAN OF THE GECMETRIC                      C
C                       DISTRIBUTION (REAL*8)                            C
C                                                                        C
C              DIST2M - INPUT MEAN OF THE NEGATIVE BINOMIAL              C
C                       DISTRIBUTION (REAL*8)                            C
C                                                                        C
C              DIST2K - INPUT PARAMETER K IN THE NEGATIVE                C
C                       BINOMIAL DISTRIBUTION (REAL*8)                   C
C                                                                        C
C              A1 - INPUT PARAMETER IN THE TRANSITION                    C
C                       MATRIX (REAL*8)                                  C
C                                                                        C
C              A2 - INPUT PARAMETER IN THE TRANSITION                    C
C                       MATRIX (REAL*8)                                  C
C                                                                        C
C          2.  REQUIRED SUBPROGRAMS:                                     C
C              INTEGER FUNCTION GEOMET                                   C
C              INTEGER FUNCTION NEGBIN                                   C
C              SUBROUTINE OVFLOW (NPS ROUTINE)                           C
C              SUBROUTINE SRAND  (NPS ROUTINE)                           C
C              SUBROUTINE SNORM  (NPS ROUTINE)                           C
C              SUBROUTINE SEXPON (NPS ROUTINE)                           C
C     ------------------------------------------------------------      C
C                                                                        C
      SUBROUTINE MODEL(X,SIZE,IX,DIST1M,DIST2M,DIST2K,A1,A2)
      IMPLICIT REAL*8 (A-H,K,O-Z)
      REAL*4 R
      INTEGER*4 GEOMET,STATE,SIZE
      DIMENSION ALPHA(2),X(SIZE)
      COMMON ETA,K,PARAMG
      CALL OVFLOW
C
C     PARAMETERS FOR NEGBIN
      K=DIST2K
      ETA=K/(DIST2M-1.0D0)
C
C     PARAMETER FOR GEOMET
      PARAMG=DLOG(1.0D0-1.0D0/DIST1M)
C
C     STATE ONE TRANSITION PROBABILITIES FOR MATRIX
      ALPHA(1)=A1
      ALPHA(2)=1.0D0-A2
C
C     SELECT INITIAL STATE FROM STEADY-STATE PROBABILITIES
      STATE=1
      CALL SRAND(IX,R,1)
```

```
      IF(R.LE.((1.0D0-ALPHA(1))/(1.0D0-ALPHA(1)+ALPHA(2))
     C)) STATE=2
C
C     COMPUTE 'SIZE' INTEREVENT TIMES
      DO 2 I=1,SIZE
C
C     ENTER MATRIX AND DETERMINE TYPE OF NEXT INTERVAL
      CALL SRAND(IX,R,1)
      IF(DBLE(R).LE.ALPHA(STATE)) GO TO 1
C
C     PICK TYPE TWO VARIATE
      STATE=2
      X(I)=DFLOAT(NEGBIN(IX))
      GO TO 2
C
C     PICK TYPE ONE VARIATE
    1 STATE=1
      X(I)=DFLOAT(GEOMET(IX))
    2 CONTINUE
      RETURN
      END
```

```
C                      INTEGER FUNCTION GEOMET                        C
C                                                                     C
C     A.  PURPOSE:                                                    C
C         THIS FUNCTION GENERATES VARIATES FROM THE                   C
C         GEOMETRIC DISTRIBUTION WHICH IS OF THE FORM                 C
C                                                                     C
C                                   X-1                               C
C                 M(X)=(1-P)*(P)         ;0<P<1;X=1,2,...             C
C                                                                     C
C     B.  USAGE:                                                      C
C         THIS FUNCTION WAS WRITTEN TO BE USED WITH                   C
C         SUBROUTINE MODEL.                                           C
C                                                                     C
C         P=1-1/(1-DIST1M)                                            C
C                                                                     C
C         PARAMG=DLOG(P)                                              C
C                                                                     C
C     ----------------------------------------------------------------C
C                                                                     C
      INTEGER FUNCTION GEOMET*4 (IX)
      IMPLICIT REAL*8 (A-H,K,O-Z)
      REAL*4 R
      COMMON ETA,K,PARAMG
C
C     CALCULATE VARIATE
      CALL SRAND(IX,R,1)
      RATIO = DLOG(DBLE(R))/PARAMG
      IRATIO = IDINT(RATIO)
C
C     ROUND UP IF NON-INTEGER
      IF (RATIO-DFLOAT(IRATIO).GT.0.0D0) IRATIO=IRATIO+1
C
C     RETURN VARIATE
      GEOMET=IRATIO
      RETURN
      END
```

```
C                      INTEGER FUNCTION NEGBIN                             C
C                                                                          C
C     A.  PURPOSE:                                                         C
C         THIS FUNCTION GENERATES VARIATES FROM THE NEGATIVE              C
C         BINOMIAL DISTRIBUTION WHICH IS OF THE FORM                       C
C                                                                          C
C                   (K-X-2)            K     X-1                           C
C         M(X)=     (      ) *(1-P) *(P)          ;K>0;0<P<1;X=1,2,...     C
C                   ( X-1  )                                               C
C                                                                          C
C     B.  USAGE:                                                           C
C         THIS FUNCTION WAS WRITTEN TO BE USED WITH                        C
C         SUBROUTINE MODEL.                                                C
C                                                                          C
C         P=1/(1+ETA)                                                      C
C                                                                          C
C     ---------------------------------------------------------------     C
C                                                                          C
        INTEGER FUNCTION NEGBIN*4 (IX)
        IMPLICIT REAL*8 (A-H,K,O-Z)
        REAL*4 U(100),E,NORM
        COMMON ETA,K,PARAMG
        GAMMAD=0.0D0
        GAMMAI=0.0D0
C
C     DETERMINE INTEGER AND DECIMAL PARTS OF K
        IK=IDINT(K)
        DK=K-DFLOAT(IK)
C
C     CALCULATE, IF REQUIRED, GAMMA IK VARIATE
C     FROM SUM OF IK UNIT EXPONENTIAL VARIATES
        IF(K.LT.1.0D0) GO TO 9
        ET=0.0D0
        DO 8 M=1,IK
        CALL SEXPON (IX,E,1)
        ET=ET+DBLE(E)
      8 CONTINUE
        GAMMAI=ET/ETA
C
C     CALCULATE, IF REQUIRED, GAMMA DK VARIATE
C     USING JOHNK'S METHOD
      9 IF(DK.LE.0.0D0) GO TO 11
     10 CALL SRAND (IX,U,2)
        UK1=DBLE(U(1))**(1.0D0/DK)
        UK2=DBLE(U(2))**(1.0D0/(1.0D0-DK))
        ZZ=UK1+UK2
        IF(ZZ.GE.1.0D0) GO TO 10
        CALL SEXPON (IX,E,1)
        GAMMAD=(UK1/ZZ)*DBLE(E)/ETA
C
C     TOTAL GAMMA VARIATE
     11 GAMMA=GAMMAD+GAMMAI
        IGAMMA=IDINT(GAMMA)
        IF(IGAMMA.GE.100) GO TO 50
C
C     CALCULATE POISSON VARIATE, DIRECTLY
        NN=0
        UT=1.0D0
        EGAMMA=DEXP(-GAMMA)
     20 CALL SRAND (IX,U,100)
        DO 30 M=1,100
        UT=UT*DBLE(U(M))
        IF(UT.LE.EGAMMA) GO TO 40
     30 CONTINUE
        NN=NN+100
        GO TO 20
     40 N=M-1+NN
        GO TO 60
C
C     CALCULATE POISSON VARIATE, USING NORMAL APPROXIMATION
     50 CALL SNORM (IX,NORM,1)
```

```
      N=IDINT(GAMMA-0.25D0+(NORM/2.0D0)**2+NORM*
     C(DSQRT(GAMMA+0.125D0)))
C
C     RETURN VARIATE
   60 NEGBIN=N+1
      RETURN
      END




C                        SUBROUTINE KSTEST                                 C
C                                                                          C
C     A.   PURPOSE:                                                        C
C          THIS SUBROUTINE CALCULATES THE PERIODOGRAM OF                   C
C          INTERVAL DATA, ESTIMATES THE POWER SPECTRAL                     C
C          DENSITY PSD AND TESTS THE FIT OF THE PERICDOGRAM                C
C          TO THE PSD.                                                     C
C                                                                          C
C     B.   USAGE:                                                          C
C          1.   ARGUMENTS:                                                 C
C                                                                          C
C               X - INPUT VECTOR OF INTERVAL DATA (REAL*8)                 C
C                                                                          C
C               IVEC - OUTPUT VECTOR OF DESIRED PERICDOGRAM                C
C                      POINTS (REAL*8)                                     C
C                                                                          C
C               MEAN - OUTPUT MEAN OF INTERVAL DATA (REAL*8)               C
C                                                                          C
C               VARIAN - OUTPUT VARIANCE OF INTERVALS (REAL*8)             C
C                                                                          C
C               XKSDN - OUTPUT KOLMOGOROV-SMIRNOV STATISTIC                C
C                       FOR TEST OF FIT (REAL*8)                           C
C                                                                          C
C               SIZE - INPUT LENGTH OF VECTOR X (INTEGER)                  C
C                                                                          C
C               MHAT - OUTPUT ESTIMATE OF COVARIANCE PARAMETER             C
C                      M (REAL*8)                                          C
C                                                                          C
C               BHAT - PUTPUT ESTIMATE OF COVARIANCE PARAMETER             C
C                      BETA (REAL*8)                                       C
C                                                                          C
C          2.   REQUIRED SUBROUTINE:  FFTR (IMSL ROUTINE)                  C
C                                                                          C
C          3.   CAUTION:  VECTOR X OF INTERVALS IS DESTROYED               C
C                         BY FFTR.                                         C
C                                                                          C
C     ----------------------------------------------------------------    C
C
      SUBROUTINE KSTEST(X,IVEC,MEAN,VARIAN,KSDN,SIZE,MHAT,
     CBHAT)
      IMPLICIT REAL*8 (A-H,O-Z)
      COMPLEX*16 GAMN
      REAL*8 IVEC,MEAN,MHAT,IB,IC,KSDN,KL,KU
      INTEGER*4 SIZE,SS,HE
      DIMENSION X(820),S(409),IWK(2495),IVEC(409)
      DATA PI/3.14159265D0/
      GAMMA1=0.0D0
      GAMMA2=0.0D0
      KSDN=0.0D0
      MEAN=0.0D0
      VARIAN=0.0D0
      NN = SIZE - IDINT(DFLOAT(SIZE)/2) - 1
      SS= NN -1
      HE = SIZE -1
      JE = SIZE - 2
C
C     CALCULATE MEAN AND VARIANCE OF INTERVALS
      DO 10 J=1,SIZE
      MEAN = MEAN + X(J)
```

```
      VARIAN = VARIAN + X(J)**2
   10 CONTINUE
      MEAN = MEAN/SIZE
      VARIAN = ( VARIAN - SIZE * MEAN**2) / (SIZE - 1)
C
C   CALCULATE ESTIMATES OF M AND BETA
      DO 40 J=1,HE
      GAMMA1 = GAMMA1 + (X(J+1)-MEAN) * (X(J)-MEAN)
   40 CONTINUE
      DO 50 J=1,JE
      GAMMA2 = GAMMA2 + (X(J+2)-MEAN) * (X(J)-MEAN)
   50 CONTINUE
      GAMMA1=GAMMA1/SIZE
      GAMMA2=GAMMA2/SIZE
      BHAT = GAMMA2 / GAMMA1
      MHAT = GAMMA1**2 / GAMMA2
C
C   COMPUTE FINITE FOURIER TRANSFORM OF INTERVAL DATA
      CALL FFTR(X,GAMN,SIZE,IWK)
C
C   CALCULATE PERIODOGRAM
      DO 20 J=3,SIZE,2
      I=(J-1)/2
      IVEC(I)=(X(J)**2 + X(J+1)**2)/(PI * SIZE)
   20 CONTINUE
C
C  TEST PERIODOGRAM FIT TO ESTIMATED POWER SPECTRAL DENSITY
      OMEGA = DCOS(2 * PI /SIZE)
      PHAT=(VARIAN+2*MHAT*BHAT*(OMEGA-BHAT)/(1+BHAT**2-2*
     CBHAT*OMEGA))/PI
      S(1) = IVEC(1)/PHAT
      DO 60 J=2,NN
      OMEGA = DCOS(2 * PI * J / SIZE)
      PHAT=(VARIAN+2*MHAT*BHAT*(OMEGA-BHAT)/(1+BHAT**2-2*
     CBHAT*OMEGA))/PI
      S(J)= S(J-1) + IVEC(J)/ PHAT
   60 CONTINUE
      DO 70 J=1,SS
      S(J) = S(J)/S(NN)
      KL= DABS(S(J)-(DFLOAT(J-1)/NN))
      KU= DABS(S(J)-(DFLOAT(J)/NN))
      KSDN = DMAX1(KSDN,KL,KU)
   70 CONTINUE
      KSDN=KSDN*DSQRT(DFLOAT(NN))
   80 RETURN
      END




C                          SUBROUTINE ESTIM8                            C
C                                                                       C
C   A.   PURPOSE:                                                       C
C        THIS SUBROUTINE USES THE METHOD OF SCORING TO                  C
C        STABILIZE ESTIMATES OF THE PARAMETERS FOR A                    C
C        UNIVARIATE TWO-STATE SEMI-MARKOV MODEL.                        C
C                                                                       C
C   B.   USAGE:                                                         C
C        1.  ARGUMENTS:                                                 C
C                                                                       C
C            M1,M2,S2,A1,A2 - INPUT INITIAL ESTIMATES FOR               C
C                             THE MEAN OF TYPE 1 INTERVALS,             C
C                             MEAN AND STANDARD DEVIATION               C
C                             OF TYPE 2 INTERVALS AND THE               C
C                             TRANSITION PROBABILITIES -                C
C                             ALL REAL*8                                C
C                                                                       C
C            SIZE - INPUT NUMBER OF INTERVAL DATA POINTS                C
C                   (INTEGER)                                           C
C                                                                       C
```

```
C           MEAN - INPUT MEAN OF INTERVAL DATA (REAL*8)      C
C                                                            C
C           IVEC - INPUT VECTOR OF PERIODOGRAM POINTS        C
C                  (REAL*8)                                  C
C                                                            C
C           ITER8 - INPUT NUMBER OF ITERATIONS DESIRED       C
C                   PRIOR TO TERMINATION IF NO CONVERGENCE   C
C                   (INTEGER)                                C
C                                                            C
C           CONVRG - CONVERGENCE CRITERION FOR TERMINATION   C
C                    (REAL*8)  10.0E-10 RECOMMENDED          C
C                                                            C
C       2.  SUBROUTINES REQUIRED:                            C
C          DMINV (IBM ROUTINE)                               C
C          DTERM (IBM ROUTINE)                               C
C                                                            C
C       3.  IF NO SCALING IS DESIRED, SET MEAN=1.0D0 .       C
C                                                            C
C    ----------------------------------------------------    C
C
      SUBROUTINE ESTIM8(M1,M2,S2,A1,A2,SIZE,MEAN,IVEC,ITER8,
     CCONVRG)
      IMPLICIT REAL*8 (A-Z)
      INTEGER*4 SIZE,NN,J,NE,JE,HE,I,L,M,N,ITE,ITER8
      DIMENSION AVEC(5),AMAT(5,5),MVEC(5),MMAT(5,5),BVEC(5)
      DIMENSION BMAT(5,5),L(5),LVEC(5),LMAT(5,5),DELT(5)
      DIMENSION LLMAT(5,5),M(5),PVEC(5),IVEC(409)
      DATA PI/3.141592654D0/,N/5/
      ITE=0
      NN=SIZE-IDINT(DFLOAT(SIZE)/2)-1
C
C    ZERO-OUT VECTORS AND MATRICES
   90 DO 100 J=1,5
      AVEC(J) = 0.0D0
      BVEC(J) = 0.0D0
      LVEC(J) = 0.0D0
      MVEC(J) = 0.0D0
      DO 110 I=1,5
      AMAT(I,J) = 0.0D0
      BMAT(I,J) = 0.0D0
      LMAT(I,J) = 0.0D0
      LLMAT(I,J)= 0.0D0
      MMAT(I,J) = 0.0D0
  110 CONTINUE
  100 CONTINUE
C
C    COUNT ITERATIONS
      ITE = ITE + 1
      WRITE (6,345) ITE
C
C    CALCULATE ELEMENTS OF EQUATIONS
      CC1 = 1 - A1
      CC2 = 1 - A2
      CC3 = CC1 + CC2
      CC4 = (M1**2) - M1
      CC5 = M1 - M2
      CC6 = CC1 * CC2
      AE = ((CC2 * CC4) + (CC1 * S2**2))/CC3
      AVEC(1)= MEAN *((1-2*M1)*(-CC2))/CC3
      AVEC(3)= MEAN * (2* S2 * CC1)/CC3
      AVEC(4)= (CC2 *(CC4 - S2**2))/ (CC3**2)
      AVEC(5)=((-CC1)*(CC4 -S2**2))/ (CC3**2)
      AMAT(1,1)= MEAN *(2 * CC2)/CC3
      AMAT(1,4)= AVEC(1) / (CC3 * MEAN)
      AMAT(1,5)= (CC1 *(1-2*M1))/(CC3**2)
      AMAT(3,3)= MEAN *(2* CC1)/ CC3
      AMAT(3,4)= (2 * S2 * (-CC2)) / CC3**2
      AMAT(3,5)= AVEC(3) / (CC3 * MEAN)
      AMAT(4,4)= (2 * AVEC(4)) / CC3
      AMAT(4,5)= ((A1-A2) * (CC4 - S2**2))/ CC3**3
      AMAT(5,5)= (2 * AVEC(5))/CC3
      ME = (CC6 * CC5**2)/CC3**2
```

```
      MVEC(1)= MEAN * (2 * CC5 * CC6) / CC3**2
      MVEC(2)= -MVEC(1)
      MVEC(4)=(CC2*(A2-A1)*(CC5**2))/CC3**3
      MVEC(5)=(CC1*(A1-A2)*(CC5**2))/CC3**3
      MMAT(1,1)=MEAN*(2*CC6)/CC3**2
      MMAT(1,2)=-MMAT(1,1)
      MMAT(1,4)=(2*CC2*(A2-A1)*CC5)/CC3**3
      MMAT(2,4)=-MMAT(1,4)
      MMAT(1,5)=(2*CC1*(A1-A2)*CC5)/CC3**3
      MMAT(2,5)=-MMAT(1,5)
      MMAT(2,2)=MEAN*(2*CC6)/CC3**2
      MMAT(4,4)=(((6*A2)+((-4)*A2**2)+(2*A1*A2)+((-2)*A1)-2)
     C*(CC5**2))/CC3**4
      MMAT(4,5)=(((2*(1-A1-A2+2*A1*A2))-A1**2-A2**2)*CC5**2)
     C/CC3**4
      MMAT(5,5)=(((6*A1)-4*A1**2+2*A1*A2-2*A2-2)*(CC5**2))
     C/CC3**4
      DO 120 J=1,3
      DO 130 I=1,5
      AMAT(J,I)=AMAT(J,I)*MEAN
      MMAT(J,I)=MMAT(J,I)*MEAN
  130 CONTINUE
  120 CONTINUE
      BETA=A1+A2-1.0D0
      LIKE=0.0D0
      DO 140 NE=1,NN
      COSINE = DCOS( 2 * PI * NE/ SIZE)
      CC7 = (1 + BETA**2) - 2 * BETA * COSINE
      BE = 1 + ( 2 *((BETA * COSINE)-BETA**2))/CC7
      BVEC(4)= 2*(((1+BETA**2)*COSINE)-2*BETA)/CC7**2
      BVEC(5)=BVEC(4)
      BMAT(4,4)=2*(6*BETA**2-2*COSINE*BETA**3+4*COSINE**2-6*
     CBETA*COSINE-2)/CC7**3
      BMAT(4,5)=BMAT(4,4)
      BMAT(5,5)=BMAT(4,4)
C
C    CALCULATE ESTIMATE OF POWER SPECTRAL DENSITY
      PE= (AE + ME *BE)/ PI
      CC8 =(IVEC(NE) -PE )/ PE**2
      CC9 = (PE - 2 * IVEC(NE)) / PE**3
C
C    CALCULATE FIRST-ORDER PARTIALS OF PSD
      DO 150 J=1,5
      PVEC(J) = (AVEC(J) + MVEC(J)*BE + BVEC(J)* ME) / PI
  150 CONTINUE
C
C    CALCULATE LIKELIHOOD FUNCTION VALUE
      LIKE = LIKE - (IVEC(NE) / PE ) - DLOG(PE)
C
C    CALCULATE SCORES
      DO 160 JE=1,5
      LVEC(JE)=LVEC(JE) + CC8 * PVEC(JE)
      DO 170 HE=JE,5
C
C    CALCULATE SECOND-ORDER PARTIALS OF PSD
      PJH=AMAT(JE,HE) + MMAT(JE,HE) * BE + MVEC(JE) *
     CBVEC(HE)+MVEC(HE)*BVEC(JE)+BMAT(JE,HE)*ME
C
C    CALCULATE SECOND-ORDER PARTIALS OF LIKELIHOOD FUNCTION
      LMAT(JE,HE)=LMAT(JE,HE)-CC9*PVEC(JE)*PVEC(HE)-CC8*
     CPJH/PI
C
C    CALCULATE EXPECTED VALUE OF LMAT
      LLMAT(JE,HE)=LLMAT(JE,HE)+PVEC(JE)*PVEC(HE)/PE**2
  170 CONTINUE
  160 CONTINUE
  140 CONTINUE
C
C    FILL IN LOWER TRIANGLE OF MATRICES
      DO 180 JE=1,4
      J=JE + 1
      DO 190 HE=J,5
```

```
      LMAT(HE,JE)=LMAT(JE,HE)
      LLMAT(HE,JE)=LLMAT(JE,HE)
  190 CONTINUE
  180 CONTINUE
      WRITE(6,360)
      DO 200 I=1,5
      WRITE(6,370) (LMAT(I,J),J=1,5)
      DO 185 J=1,5
      AMAT(I,J)=-LMAT(I,J)
  185 CONTINUE
  200 CONTINUE
C
C   CALCULATE INVERSE AND DETERMINANT OF LMAT
      CALL DMINV(LMAT,N,D,L,M)
      WRITE(6,380) D
      WRITE(6,390)
      DO 210 I=1,5
      WRITE(6,370) (LMAT(I,J),J=1,5)
  210 CONTINUE
C
C   CALCULATE ADDITIVE INCREMENTS TO ESTIMATES
      DO 240 I=1,5
      DELT(I)= 0.0D0
      DO 250 J=1,5
      DELT(I)= DELT(I) + LMAT(I,J) * LVEC(J)
  250 CONTINUE
  240 CONTINUE
C
C   INCREMENT ESTIMATES
      M1= M1 + DELT(1) * MEAN
      M2= M2 + DELT(2) * MEAN
      S2= S2 + DELT(3) * MEAN
      A1= A1 + DELT(4)
      A2= A2 + DELT(5)
      WRITE(6,430) LIKE
      WRITE(6,440) (LVEC(J),J=1,5)
      WRITE(6,450) M1,M2,S2,A1,A2
C
C   TEST FOR CONVERGENCE
      IF(DMAX1(DABS(DELT(1)),DABS(DELT(2)),DABS(DELT(3)),
     CDABS(DELT(4)),DABS(DELT(5))).LE.CONVRG) GO TO 260
      IF(DMAX1(DABS(LVEC(1)),DABS(LVEC(2)),DABS(LVEC(3)),
     CDABS(LVEC(4)),DABS(LVEC(5))).LE.CONVRG) GO TO 260
C
C   TEST FOR NUMBER OF ITERATIONS
  255 IF(ITE.LT.ITER8) GO TO 270
  280 WRITE (6,455) CONVRG
      GO TO 265
  270 WRITE(6,460)
C
C   CONTINUE IF NECESSARY
      GO TO 90
  260 WRITE(6,470)
C
C   TEST LMAT FOR NEGATIVE DEFINITENESS
  265 DO 275 I=1,5
      DO 267 JE=1,5
      DO 266 HE=1,5
      BMAT(JE,HE)=AMAT(JE,HE)
  266 CONTINUE
  267 CONTINUE
      CALL DTERM(I,BMAT,D,N)
      AVEC(I)=D
  275 CONTINUE
      IF(AVEC(1).LT.0.0D0.AND.AVEC(2).GT.0.0D0.AND.AVEC(3)
     C.LT.0.0D0.AND.AVEC(4).GT.0.0D0.AND.AVEC(5).LT.0.0D0)
     C GO TO 290
  285 WRITE(6,480)
      DO 295 I=1,5
      WRITE(6,490) I,AVEC(I)
  295 CONTINUE
      GO TO 600
```

```
  290 WRITE(6,500)
  600 WRITE(6,400)
      DO 220 I=1,5
      WRITE(6,370) (LLMAT(I,J),J=1,5)
  220 CONTINUE
C
C    INVERT EXPECTED VALUE MATRIX
      CALL DMINV(LLMAT,N,D,L,M)
      WRITE(6,410) D
      WRITE(6,420)
C
C    TEST INVERSE FOR POSITIVE DEFINITENESS
      DO 230 I=1,5
      WRITE(6,370) (LLMAT(I,J),J=1,5)
      DO 225 J=1,5
      AMAT(I,J)=LLMAT(I,J)
  225 CONTINUE
  230 CONTINUE
      DO 610 I=1,5
      DO 620 JE=1,5
      DO 630 HE=1,5
      BMAT(JE,HE)=AMAT(JE,HE)
  630 CONTINUE
  620 CONTINUE
      CALL DTERM(I,BMAT,D,N)
      AVEC(I)=D
  610 CONTINUE
      IF(DMIN1(AVEC(1),AVEC(2),AVEC(3),AVEC(4),AVEC(5)).GT.
     C0.0D0) GO TO 660
  640 WRITE(6,510)
      DO 650 I=1,5
      WRITE(6,520) I,AVEC(I)
  650 CONTINUE
      RETURN
  660 WRITE(6,530)
      RETURN
  345 FORMAT(1H1,//////////,' ITERATION NUMBER',I4)
  360 FORMAT(//,' NEGATIVE MATRIX OF SECOND PARTIALS',/)
  370 FORMAT(/,5D20.10)
  380 FORMAT(//,' DETERMINANT OF MATRIX',//,D20.10)
  390 FORMAT(//,' INVERSE MATRIX',/)
  400 FORMAT(//,' INFORMATION MATRIX',/)
  410 FORMAT(//,' DETERMINANT OF INFORMATION MATRIX',//,
     CD20.10)
  420 FORMAT(//,' INVERSE INFORMATION MATRIX',/)
  430 FORMAT(//,' LOG LIKELIHOOD FUNCTION VALUE =',D20.10)
  440 FORMAT(//,' SCORES =',5D20.10)
  450 FORMAT(//,' PARAMETERS =',5D20.10)
  455 FORMAT(//,' NUMBER OF ITERATIONS REQUESTED WAS',
     C' REACHED AND',//,' CONVERGENCE CRITERION ( ',
     CD18.10,' ) WAS NOT ACHIEVED.',//,' RUN TERMINATED')
  460 FORMAT(//,' CONTINUING')
  470 FORMAT(//,' COVERGENCE')
  480 FORMAT(1H1,/////,' MATRIX OF SECOND PARTIALS IS NOT ',
     C'NEGATIVE DEFINITE',/)
  490 FORMAT(/,' DETERMINANT OF PRINCIPAL MINOR SUBMATRIX ',
     C'OF ORDER',I2,' OF MATRIX OF SECOND PARTIALS = ',
     CD20.10)
  500 FORMAT(1H1,/////,' MATRIX OF SECOND PARTIALS IS ',
     C'NEGATIVE DEFINITE')
  510 FORMAT(///,' INVERSE INFORMATION MATRIX IS NOT ',
     C'POSITIVE DEFINITE',/)
  520 FORMAT(/,' DETERMINANT OF PRINCIPAL MINOR SUBMATRIX ',
     C'OF ORDER',I2,' OF INVERSE INFORMATION MATRIX = ',
     CD20.10)
  530 FORMAT(///,' INVERSE INFORMATION MATRIX IS POSITIVE ',
     C'DEFINITE')
      END
```

LIST OF REFERENCES

Cooley, J. W., Lewis, P. A. W., and Welch, P. D., "The Application of the Fast Fourier Transform Algorithm to the Estimation of Spectra and Cross-Spectra," J. Sound Vib., V. 12, p. 339-352, 1970.

Cooley, J. W., Lewis, P. A. W., and Welch, P. D., "The Fast Fourier Transform and its Application to Time Series Analysis," Chap. 13 of Statistical Methods for Digital Computers, Edited by Enslein, K., Ralston, A., and Wilf, H. S., to be published 1974.

Cox, D. R. and Lewis, P. A. W., The Statistical Analysis of Series of Events, Methuen, 1966.

International Mathematical and Statistical Libraries, Inc., Library 1, 3d ed., V. I, IMSL, INC., Revised 1 July 1973.

Johnk, M. D., "Erzeugung von Betaverteilten und Gammaverteilten Zufallszahlen," Metrika, V. 8, p. 5-12, 1964.

Jolley, L. B. W., Summation of Series, 2d ed., Dover Publications, 1961.

Naval Postgraduate School NPS 55LW73061A, Naval Postgraduate School Random Number Generator Package LLRANDOM, by G. P. Learmonth and P. A. W. Lewis, June 1973.

Lewis, P. A. W. and Shedler, G. S., "Empirically Derived Micromodels for Sequences of Page Exceptions," IBM Journal of Research and Development, V. 17, p. 86-100, March 1973.

Lilliefors, H. W., "On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown," J. of the Amer. Stat. Assoc., V. 62, p. 399-402, June 1967.

Lilliefors, H. W., "On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown," J. of the Amer. Stat. Assoc., V. 64, p. 387-389, March 1969.

Olshen, R. A., "Asymtotic Properties of the Periodogram of a Discrete Stationary Process," J. Appl. Prob., V. 4, p. 508-528, 1967.

Rao, C. R., Linear Statistical Inference and its Applications, p. 302-305, Wiley, 1965.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Chief of Naval Personnel Pers 11b Department of the Navy Washington, D. C. 20370	1
3.	Department of Operations Research and Administrative Sciences, Code 55 Naval Postgraduate School Monterey, California 93940	2
4.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
5.	Professor P. A. W. Lewis, Code 55LW Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	4
6.	Asst Professor M. U. Thomas, Code 55TO Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
7.	Gerald S. Shedler IBM Research Labs, K08/282 Monterey and Cottle Rds. San Jose, California 95114	1
8.	Dr. Jean Paul Jacob IBM Research Labs Monterey and Cottle Rds. San Jose, California 95114	1
9.	Dr. E. Altman IBM Research Labs Monterey and Cottle Rds. San Jose, California 95114	1
10.	LT James Leroy Hornback, USN 8412 Oswego Street Sunland, California 91040	1

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Parameter Estimation for a Two-State Semi-Markov Model of a Univariate Point Process		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis March 1974
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) James Leroy Hornback		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE March 1974
		13. NUMBER OF PAGES 41
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Naval Postgraduate School Monterey, California 93940		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

The author is indebted to IBM San Jose Research Labs for use of their facilities during his experience tour. No IBM confidential material or information was utilized in this thesis.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Univariate Point Process, Two-State Semi-Markov Model, Parameter Estimation, Multiprogrammed computers, Page exceptions, Page exception processes, Power Spectral Density, Periodogram, Page Reference Patterns, Demand Paged Computer System.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Using the convenient second-order interval properties of a two-state semi-Markov model for a univariate point process, an automated technique for the estimation of the parameters in the model was researched and discussed. The power spectral density of intervals was estimated by the periodogram and a Kolmogorov-Smirnov test of fit was conducted. The asymtotic exponential distribution and independence of the periodogram points were

Block 20 - Abstract (Cont.)

used to calculate an approximate likelihood function. A system of equations was then formed to find the maximum likelihood estimates of the parameters. Since closed-form solutions for the estimates could not be found, an iterative method to stabilize initial guesses of the parameter values was attempted with only limited success. Results on using Kolmogorov-Smirnov type statistics and the spectrum of intervals to test the fit of stochastic process models to data have also been obtained.

Thesis 140766
H8015 Hornback
c.1 Parameter estimation for a two-state semi-Markov model of a univariate point process.